Title
Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses.

Journal
Virology, 183(2)

ISSN
1096-0341

Authors
Yokomori, K
Banner, LR
Lai, MM

Publication Date
1991-08-01

DOI
10.1016/0042-6822(91)90994-m

Peer reviewed

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 183, 647–657 (1991)

Heterogeneity of Gene Expression of the Hemagglutinin-Esterase (HE) Protein of Murine Coronaviruses

KYOKO YOKOMORI, LISA R. BANNER, AND MICHAEL M. C. LAI[1]

Howard Hughes Medical Institute and Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, California 90033

Received January 28, 1991; accepted April 25, 1991

The hemagglutinin-esterase (HE) membrane glycoprotein is present only in some members of the coronavirus family, including some strains of mouse hepatitis virus (MHV). In the JHM strain of MHV, expression of the HE gene is variable and corresponds to the number of copies of a UCUAA pentanucleotide sequence present at the 3′-end of the leader RNA. This copy number varies among MHV strains, depending on their passage history. The JHM isolates with two copies of UCUAA in their leader RNA showed a high level of HE expression, whereas the JHM isolate with three copies had a low-level expression. In this study, the analysis of HE gene expression was extended to other MHV strains. The synthesis of HE mRNA in these viruses also correlates with the copy number of UCUAA in the leader RNA and the particular intergenic sequence preceding the HE gene. In one MHV strain, MHV-1, no detectable HE mRNA was synthesized, despite the presence of a proper transcription initiation signal. This lack of HE mRNA expression was consistent with a leader RNA containing three UCUAA copies. However, mutations and deletions within the coding region of the MHV-1 HE gene have generated a stretch of sequence which resembled the transcriptional initiation motif, and was shown to initiate the synthesis of a novel smaller mRNA. These findings strengthened the theory that interactions between leader RNA and transcriptional initiation sequences regulate MHV subgenomic mRNA transcription. Sequence analysis revealed that most MHV strains, through extensive mutations, deletions, or insertions, have lost the complete HE open reading frame, thus turning HE into a pseudogene. This high degree of variation is unusual as the other three structural proteins (spike, membrane, and nucleocapsid) are well-maintained. In contrast to bovine coronavirus, which apparently requires HE for viral replication, the HE protein in MHV may be only an accessory protein which is not necessary for viral replication. JHM and MHV-S, however, have preserved the expression of HE protein. © 1991 Academic Press, Inc.

INTRODUCTION

Coronavirus virions are enveloped particles containing three or four structural proteins: spike (S) protein, membrane (M) protein, nucleocapsid (N) protein, and hemagglutinin-esterase (HE) protein (Lai, 1990). The S protein is a glycoprotein of 180 kDa, which often is cleaved by host proteases into two 90-kDa subunits (Yoshikura and Tejima, 1981; Sturman *et al.*, 1985). This protein, which is the outermost protein on the viral envelope, interacts with the host cell receptor, causes cell–cell fusion, and elicits neutralizing antibodies (Collins *et al.*, 1982; Sturman and Holmes, 1983; Sturman *et al.*, 1985). The M protein is a transmembrane glycoprotein of 23 kDa, which appears to interact with the nucleocapsid inside the virus particle (Sturman *et al.*, 1980). The N protein is a phosphoprotein of 50–60 kDa in size, which interacts with viral RNA (Stohlman *et al.*, 1988). The HE protein is an additional glycoprotein of 65 kDa, which projects from the viral envelope (King *et al.*, 1985; Yokomori *et al.*, 1989). While the S, M, and N are present in all coronaviruses, the HE protein is detected only in some species of coronavirus, such as bovine coronavirus (BCV) (King *et al.*, 1985; Vlasak *et al.*, 1988a; Parker *et al.*, 1989), turkey coronavirus (TCV) (Dea and Tijssen, 1988), and human coronavirus (OC43) (Hogue *et al.*, 1984; Hogue and Brian, 1986), but not in avian infectious bronchitis virus (IBV) or porcine transmissible gastroenteritis virus (TGEV) (Garwes and Reynolds, 1981; Stern and Sefton, 1982).

Mouse hepatitis virus (MHV) is the prototypic member of the Coronaviridae, which contains a nonsegmented, positive-sensed genomic RNA of 31 kb in length (Pachuk *et al.*, 1989; Lee *et al.*, 1991). In infected cells, seven or eight virus-specific mRNA species are detected. The mRNAs are characterized by a nested-set structure with common 3′-terminal sequences but with unique 5′-regions (Lai *et al.*, 1981; Leibowitz *et al.*, 1981). Although structurally polycistronic, each mRNA encodes only one protein, which is translated from the unique 5′-region. For instance, the S protein is translated from mRNA 3, M protein from mRNA 6, and N protein from mRNA 7 (Leibowitz *et al.*,

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. M64313, M64314, M64315, and M64316.

[1] To whom reprint requests should be addressed.

0042-6822/91 $3.00

1982; Siddell, 1983). In addition, the mRNAs possess a common leader sequence of approximately 72 nucleotides in length, which is derived from the 5′-end of the viral genome (Spaan *et al.*, 1983; Lai *et al.*, 1984). The 3′-end of the leader RNA contains a pentanucleotide sequence of UCUAA, which is homologous to the intergenic sequences present at the beginning of each gene. The copy number of this pentanucleotide sequence was two to four, depending on the particular virus strain (Makino *et al.*, 1988), and varied after serial passage of virus in cell culture (Makino and Lai, 1989a). For example, in the neurotropic JHM strain of MHV, the wild-type virus contained three copies of UCUAA; but after 10 to 12 passages, the number of copies was reduced to two (Makino and Lai, 1989a). This change was accompanied by an alteration in the pattern of mRNA expression (Makino and Lai, 1989a). Specifically, the JHM variant with two copies of UCUAA (designated as JHM(2)) made a large amount of mRNA 2-1, while wild-type JHM with three copies (JHM(3)) made only a small amount (Makino and Lai, 1989a; Shieh *et al.*, 1989). *In vitro* translation studies revealed that HE protein was encoded by mRNA 2-1 (Shieh *et al.*, 1989). Thus, HE protein expression in JHM was controlled by the copy number of UCUAA in the leader (Makino and Lai, 1989a; Shieh *et al.*, 1989).

The HE protein possesses an esterase activity similar to the receptor-destroying enzyme (RDE) of influenza C virus (Vlasak *et al.*, 1988b), and also a hemagglutinin activity (Vlasak *et al.*, 1988b; Parker *et al.*, 1989). However, the hemagglutinin activity in JHM is very weak (Yokomori *et al.*, 1989). In order to better understand the biological significance of the HE protein, we examined the structure and expression of HE in several different strains of MHV. The results demonstrated that HE expression varied at both transcriptional and translational levels amongst MHV strains. In contrast to the conserved expression of the other three structural proteins, S, M, and N, the HE gene in most MHV strains appears to be a pseudogene. The significance of the maintenance of HE expression in some of the MHV strains will be discussed.

MATERIALS AND METHODS

Viruses and cell culture

The MHV strains used in this study were from several different sources. Plaque-cloned JHM(2) and JHM(3) viruses (Makino and Lai, 1989a; Shieh *et al.*, 1989; Yokomori *et al.*, 1989), which have two and three copies of UCUAA in the leader sequence, respectively, and strain A59 were the same as those reported in a previous study (Yokomori *et al.*, 1989). The two variants of JHM-DL (a large-plaque variant of JHM (Weiner, 1973; Stohlman *et al.*, 1982)) containing two or three copies of UCUAA were obtained by a modification of the published procedures (Makino and Lai, 1989a). Briefly, JHM-DL was passaged serially without dilution in DBT cells, a murine astrocytoma cell line (Hirano *et al.*, 1974). Virus harvested at the end of the fifth passage was passaged one additional time in DBT cells at low multiplicity of infection (m.o.i.) to eliminate possible DI particles. Virus variants were plaque-purified, and the number of UCUAA copies at the 3′-end of the leader was determined by primer extension (Makino *et al.*, 1988). Viruses with either two or three copies of UCUAA in the leader were designated JHM-DL(2) and JHM-DL(3), respectively. MHV-1 and MHV-S were obtained from Dr. Paul Masters of Wadsworth Center for Laboratories and Research, New York State Department of Health. MHV-2 has been described previously (Lai and Stohlman, 1981). Two isolates of MHV-3, i.e., MHV-3-L and MHV-3-Yac, were plaque-purified from stocks originally obtained from Dr. Lucie Lamontagne of the University of Montreal. Viruses were propagated in DBT cells as described previously (Makino *et al.*, 1984b). DBT cells also were used for all the *in vivo* metabolic labeling experiments.

^{35}S-labeling of intracellular proteins and preparation of cell lysates

Intracellular proteins of virus-infected DBT cells were labeled, and cell lysates made as described previously (Yokomori *et al.*, 1989). Briefly, when the cytopathic effect (CPE) in MHV-infected DBT cells reached 95%, cells were starved in methionine-free media for 30 min, and then [^{35}S]methionine (1193 Ci/mmol; ICN Translabel) was added to a final concentration of 50 μCi/ml. Cells were pulse-labeled for 20 min and then immediately placed on ice, washed with ice-cold phosphate-buffered saline (PBS), and lysed with lysis buffer (1% Triton X-100, 0.5% sodium deoxycholate, and 0.1% SDS in PBS) containing 1 mM phenylmethylsulfonyl fluoride (PMSF; Sigma). Cell lysates were passed through a 19-gauge syringe needle five times, transferred to Eppendorf tubes, and spun at 12,000*g* for 10 min at 4°. Supernatant were harvested and stored at −70° until required.

Immunoprecipitation and SDS–PAGE

Immunoprecipitation was performed as previously described (Yokomori *et al.*, 1989) using rabbit hyperimmune serum prepared against purified JHM(2) virus (Yokomori *et al.*, 1989). After immunoprecipitation, samples were denatured in 2× sample buffer (0.06 M Tris–HCl, 2% SDS, 25% glycerol, 5% 2-mercaptoethanol, and 0.1% bromphenol blue) at 37° for 20 to 30

min, and centrifuged at 12,000*g* for 5 min. The supernatants were electrophoresed on 7.5 to 15% gradient or 12.5% straight polyacrylamide gels containing 0.1% SDS.

^{32}P-labeling of intracellular RNA of virus-infected cells

DBT cells were infected with various strains of MHV at an m.o.i. of 1–5. Virus-infected cells were labeled with 200 μCi/ml of [^{32}P]orthophosphate (ICN Biochemicals) from 5.5 to 8 hr postinfection in the presence of actinomycin D (2.5 μg/ml, Sigma). RNA extraction was performed as described below.

Preparation of virion RNA and intracellular RNA

Virion RNA was purified by a modification of the procedure described by Makino *et al.* (1984a). DBT cells were infected at an m.o.i. of 1–5. Culture supernatant was harvested at 15–20 hr postinfection, and cell debris removed by centrifugation at 1500 rpm for 5 min. Virus was then precipitated by the addition of ammonium sulfate to 50% saturation, and purified by ultracentrifugation twice in discontinuous sucrose gradients (Banner *et al.*, 1990). The virus pellet was incubated at 37° for 1 hr in a solution containing 100 μg/ml of proteinase K, 100 mM Tris–HCl (pH 7.5), 12.5 mM EDTA, 150 mM NaCl, and 1% SDS, extracted twice with phenol/chloroform (1/1) and the RNA precipitated with ethanol. Virion RNA isolated was used for RNA sequencing and primer extension. Intracellular RNA was isolated essentially as described by Makino *et al.* (1984b). Briefly, infected cells were lysed at 9 hr postinfection with NTE buffer (0.1 M NaCl, 0.01 M Tris–HCl, pH 7.2, and 1 mM EDTA) containing 0.5% Nonidet P-40, and the nuclei were removed by brief centrifugation. Supernatants were treated with proteinase K and the RNA extracted as described above for virion RNA.

cDNA cloning of viral mRNAs by polymerase chain reaction (PCR)

cDNA clones corresponding to the 5′-unique coding region of mRNA 2-1 were produced using PCR as described previously (Makino *et al.*, 1988). Intracellular RNA from virus-infected cells was annealed with primer 226 (5′-CTAACACCGCTATCCGTCAT-3′), which is complementary to the 3′-conserved region of gene 2-1, and cDNA was synthesized with reverse transcriptase (Boehringer Mannheim Biochemicals). For PCR amplification of the cDNA, a second primer, 78 (5′-AGCTTTACGTACCCTCTCTACTCTAAAACTCTTGTAGTTT-3′), which is homologous to the 5′-end of the leader plus 7 additional nucleotides containing the *Sna*BI restriction site, was added (Makino and Lai, 1989b). The mixture was incubated for 25 cycles of 94° for 30 sec, 54° for 90 sec, and 72° for 3 min. PCR products representing the 5′-region of mRNA 2 (2.1 kb, consisting of 0.8 kb from gene 2 and 1.3 kb from HE gene) and 2-1 (1.3 kb) were excised from low melting agarose (Sea Plaque, FMC Bioproducts, Rockland, ME) and extracted with phenol and chloroform. The cDNA was phosphorylated and blunt-ended with polynucleotide kinase and T4 DNA polymerase (Boehringer Mannheim Biochemicals), and ligated into the *Sma*I site of the vector pTZ18U (United States Biochemical Corp).

RNA and DNA sequencing

RNA sequencing of purified virion genomic RNA was performed using modification of the dideoxyribonucleotide chain-termination method (Sanger *et al.*, 1977) as described previously (Banner *et al.*, 1990). Primer 56 (5′-CGCCGAATGGACACGTC-3′), which is complementary to nucleotides 172 to 188 from the 5′-end of genomic RNA (Makino and Lai, 1989b) was used to obtain the sequence of the leader RNA, including the UCUAA repeat region. For the HE gene sequence, DNA sequencing was carried out by dideoxyribonucleotide chain-termination method (Sanger *et al.*, 1977), using primers corresponding to different regions of the gene. Some regions also were confirmed by RNA sequencing with the same primers and the method described above.

In vitro transcription and translation

Recombinant plasmid pTZ18U(HE), derived from PCR products of various strains of MHV, were constructed such that the T7 RNA polymerase promoter preceded the leader sequence and translation initiation site of the HE gene. Plasmids were linearized by digestion with *Xba*I and transcribed *in vitro* with T7 RNA polymerase as described previously (Makino and Lai, 1989b). The RNA transcripts were translated in a mRNA-dependent rabbit reticulocyte lysate or wheat germ extract (Promega Biotec) in the presence of [^{35}S]methionine (NEN Du Pont), with or without canine pancreatic microsomal membrane (Promega Biotec). Reactions were carried out as recommended by the manufacturer.

RESULTS

Detection of intracellular virus-specific HE protein from various strains of MHV

To examine the HE expression of different MHV strains, virus-infected cells were labeled with [^{35}S]methionine, and the virus-specific proteins were im-

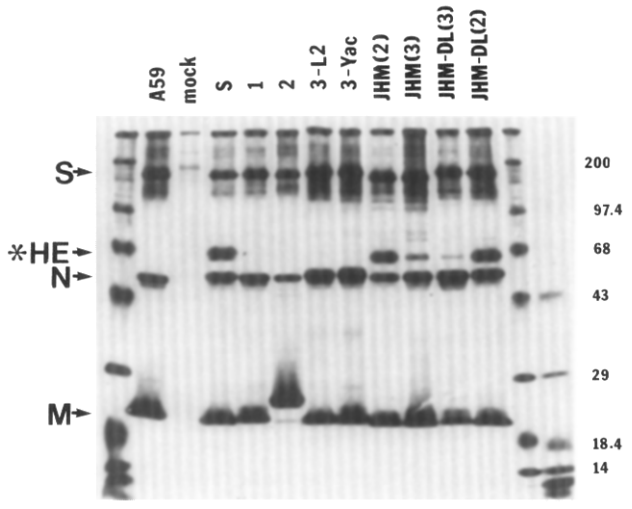

FIG. 1. SDS–PAGE analysis of intracellular viral proteins of different MHV strains. Virus-infected cells were radiolabeled with [^{35}S]methionine for 20 min at 8 to 10 hr postinfection. Cell lysates were immunoprecipitated with anti-JHM(2) rabbit serum (Yokomori *et al.*, 1989), and analyzed by electrophoresis on 7.5 to 15% gradient polyacrylamide gels. The arrows indicate viral structural proteins. Lanes S, 1, 2, and 3 represent MHV-S, MHV-1, MHV-2, and MHV-3, respectively.

munoprecipitated with polyclonal antibodies directed against JHM(2) virions. Figure 1 shows that while large amounts of HE protein could be detected in cells infected with MHV-S, JHM(2), and JHM-DL(2), only small amounts were detected in JHM(3)- and JHM-DL(3)-infected cells. This finding is consistent with our previous studies, which showed that JHM viruses with three copies of UCUAA in the leader RNA synthesized a much lower amount of HE protein than the virus with two UCUAA copies (Shieh *et al.*, 1989; Yokomori *et al.*, 1989). Since the JHM and JHM-DL variants with different copy numbers of UCUAA were isolated independently from different virus stocks, these findings further suggest that variations in the amount of HE protein are related to the copy number of UCUAA rather than some other mutation(s). Strains A59, MHV-1, MHV-2, and two different plaque isolates of MHV-3 did not synthesize HE or any other additional proteins which could be recognized by anti-JHM(2) polyclonal antibodies. These results showed that many strains of MHV do not express HE protein. In contrast to HE, the other three structural proteins (S, N, and M) were synthesized consistently by all viruses.

mRNA 2-1 expression by different strains of MHV

To determine whether the lack of synthesis of HE protein by different MHV strains was at the transcriptional or translational level, ^{32}P-labeled virus-specific mRNA from virus-infected cells was examined by agarose gel electrophoresis after glyoxal denaturation. Figure 2 shows the intracellular mRNA patterns of the various strains of MHV. Surprisingly, mRNA 2-1, which encodes HE protein, was synthesized by most of the viruses, including MHV-2, and two isolates of MHV-3, even though these strains failed to synthesize any HE-related protein (Fig. 1). Thus, the mRNA 2-1 of MHV-2 and MHV-3 was nonfunctional, or defective in translational control. Consistent with their failure to express HE protein, mRNA 2-1 was not detected in A59 or MHV-1-infected cells.

Previously we have shown that JHM(3) synthesizes less mRNA 2-1 than JHM(2) (Makino and Lai, 1989a, Shieh *et al.*, 1989). In this study, we confirmed this observation (Fig. 2) and extended it to include the large plaque variant of JHM, JHM-DL. Specifically, substantially more mRNA 2-1 was detected in cells infected with JHM-DL(2), which contains two copies of UCUAA, than JHM-DL(3), which contains three copies of UCUAA (Fig. 2). The amounts of HE protein synthesized by these viruses (Fig. 1) correlated well with the amounts of mRNA 2-1 synthesized.

Figure 2 also revealed that MHV-1, which failed to synthesize any detectable mRNA 2-1, synthesized a novel mRNA species, which was smaller than mRNA

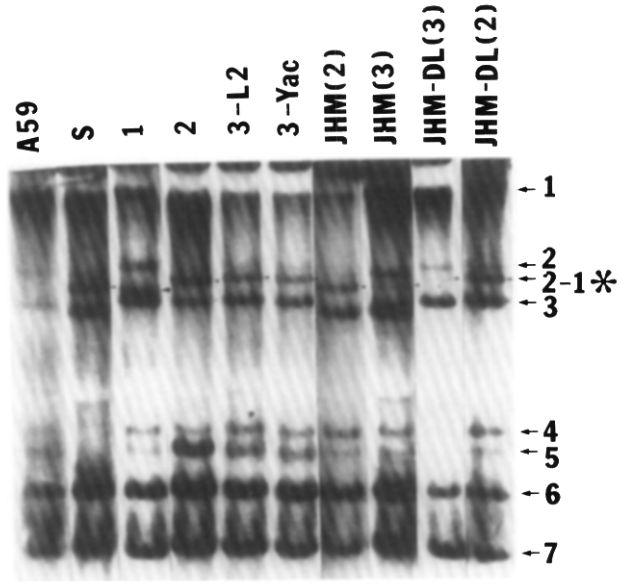

FIG. 2. Intracellular viral RNA of different MHV strains. Intracellular RNA of virus-infected cells was labeled with [^{32}P]orthophosphate for 2 hr and isolated when cytopathic effect (CPE) reached 85 to 95%. RNA was denatured with glyoxal and analyzed by electrophoresis on a 1% agarose gel. Although all viral RNAs were analyzed on one gel at the same time, different lanes from the same gel were exposed for various times to adjust the intensity of the bands. Lane designations are the same as in Fig. 1. Virus-specific mRNA species are indicated by numbers, and mRNA 2-1 is indicated by small arrowheads. The novel mRNA between mRNA 2 and mRNA 3 in MHV-1 is indicated by the larger arrowhead.

2-1 but larger than mRNA 3. Since no HE-related protein was detected in MHV-1-infected cells (Fig. 1), this mRNA also appeared to be nonfunctional.

Sequence analysis of HE gene

We wanted to understand the molecular basis of the variable HE gene expression in different MHV strains and also the discrepancy between the expression of mRNA 2-1 and synthesis of HE protein by some viruses. Therefore, the sequences of the HE gene region of different viruses were obtained by either direct RNA sequencing or DNA sequencing of PCR products. Oligonucleotides corresponding to the leader RNA and a conserved 3′-end region of the HE gene were used as primers for PCR. Due to the nested-set structure of the coronavirus mRNAs, these clones included sequences from the coding region of the HE gene, the leader-fusion site of mRNA 2-1, and its genomic intergenic sequences. Thus, the 5′-unique coding regions of both mRNA 2 and mRNA 2-1 were obtained. Sequence data from the HE genes of these viruses revealed substantial numbers of base substitutions, deletions, and insertions spread throughout the coding region of gene 2-1 in most viruses. As a result, the open reading frame (ORF) for the HE protein was truncated in A59, MHV-1, MHV-2, and both isolates of MHV-3 (Fig. 3). Specifically, in contrast to the HE genes of JHM and MHV-S, which had a coding capacity of 440 amino acids, that of A59 encoded only 14 amino acids, MHV-2, 97 amino acids, MHV-3-Y, 27 amino acids, while that of MHV-3-L was 135 amino acids. Although MHV-1 did not synthesize mRNA 2-1, its ORF was capable of encoding a protein of 268 amino acids. The usual initiating AUG of the HE gene in several viruses was also mutated or deleted (Fig. 3), thus providing the potential for additional HE gene variation in different MHV strains.

To confirm the sizes of the ORF predicted from the RNA sequence data, the HE genes of some of the viruses were examined by *in vitro* translation. Capped RNAs were transcribed *in vitro* from cDNA clones using T7 RNA polymerase, and then used for *in vitro* translation in either wheat germ extract or rabbit reticulocyte lysate system. The sizes of the ^{35}S-labeled translation products agreed with the predicted ORF (Fig. 4A): The MHV-3-L clone yielded a 14-kDa translation product, whereas the MHV-3-Y clone yielded an 4-kDa product. Furthermore, in the presence of canine pancreatic microsomal membrane, MHV-3-L RNA yielded two additional protein products (Fig. 4B). The higher band possibly represented glycosylated product, since the ORF contained one potential N-linked glycosylation site. The middle band was of the same size as the primary translation product in the absence of membrane and thus represented the native product of 14 kDa. The lower band most likely represented the core protein, from which the signal peptide had been cleaved off. These results suggest that this truncated HE protein was properly translocated and processed *in vitro.* However, these protein products were not detected in infected cells (Fig. 1).

The UCUAA copy number in the leader and intergenic sequence of the HE gene

From studies with JHM(2) and JHM(3), it was concluded that the transcription of mRNA 2-1 depends on the interaction between the repeated UCUAA sequence motif in the leader RNA and the intergenic sequences upstream of the HE gene (Makino and Lai, 1989a; Shieh *et al.,* 1989). Similar conclusions were reached with JHM-DL(2) and JHM-DL(3).

Sequence analysis of the intergenic regions of the HE gene and the leader RNA was further extended to all of the other viruses. Table 1 summarizes the intergenic sequences preceding the HE gene. All viruses examined contained a stretch of UA-rich sequence similar to that of JHM. The only exception was A59, which had an A to G substitution, and correspondingly, did not synthesize mRNA 2-1, despite the fact that it contained two copies of UCUAA in the leader RNA (Makino and Lai, 1989a; Shieh *et al.,* 1989). The intergenic sequence preceding the HE gene was slightly different from the consensus transcription initiation motif seen in most of the other MHV genes, i.e., gene 3 and gene 7 (Table 1). This finding may explain why gene 2-1 was regulated by the number of UCUAA copies in the leader RNA, while the other genes were constitutively expressed. Since MHV-1 had an identical intergenic sequence preceding the HE gene (Table 1), the failure of this virus to transcribe mRNA 2-1 must have been due to the presence of three copies of UCUAA at the 3′-end of leader sequence, or some other as yet undetermined reason.

The number of copies of UCUAA in the leader RNA of the various MHV strains was determined by primer extension and direct RNA sequencing of genomic RNA. The results showed that MHV-S, MHV-2, MHV-3-L, and MHV-3-Y had two copies while MHV-1 had three (Fig. 5 and data not shown). As MHV-S, MHV-2 and both MHV-3 isolates transcribed mRNA 2-1, while MHV-1 did not (Fig. 2), these results further supported the correlation between the UCUAA copy number and expression of the HE gene. In each case, those viruses with three repeats did not transcribe mRNA 2-1 efficiently.

As described above, MHV-1 did not synthesize mRNA 2-1, but instead, made a novel mRNA smaller

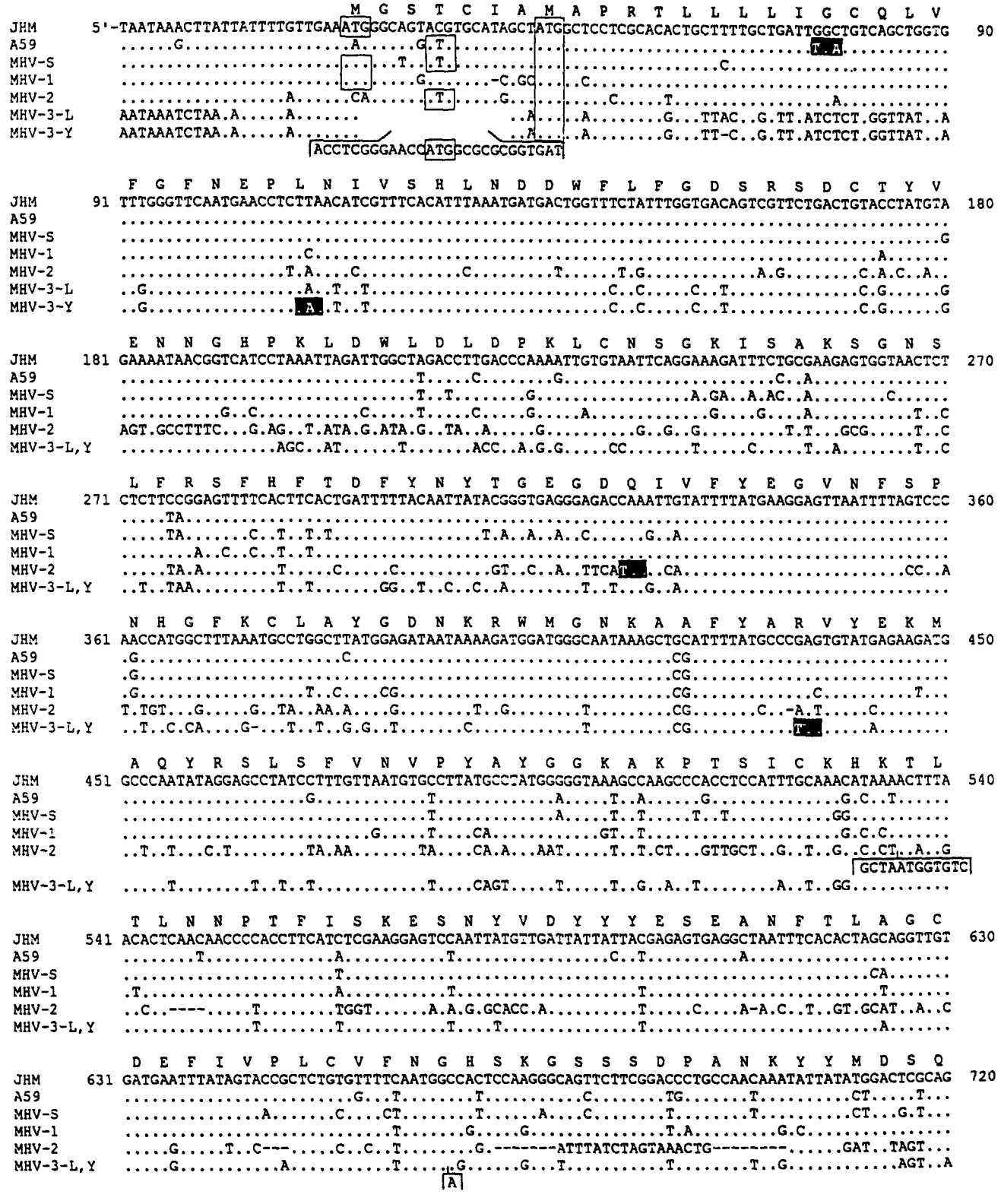

FIG. 3. Sequence of the HE gene starting from the upstream intergenic sites. JHM and A59 sequences have been published previously (Luytjes *et al.*, 1988; Shieh *et al.*, 1989). Amino acid sequences are shown above the nucleotide sequence and represent that of JHM (Shieh *et al.*, 1989). Open boxes indicate AUG initiation codons and black boxes indicate termination codons of ORFs. Dashed lines represent deletions. Insertions are shown underneath with brackets. The underlined sequence beginning at the position 797 of MHV-1 indicates the intergenic site for the novel mRNA. The double-underlined sequence at the 3'-end of the gene corresponds to the primer used for PCR. The sequences for MHV-1, -2, -3, and -4 have been deposited with GenBank. Their accession numbers are M-64313, M-64314, M-64315, and M-64316, respectively.

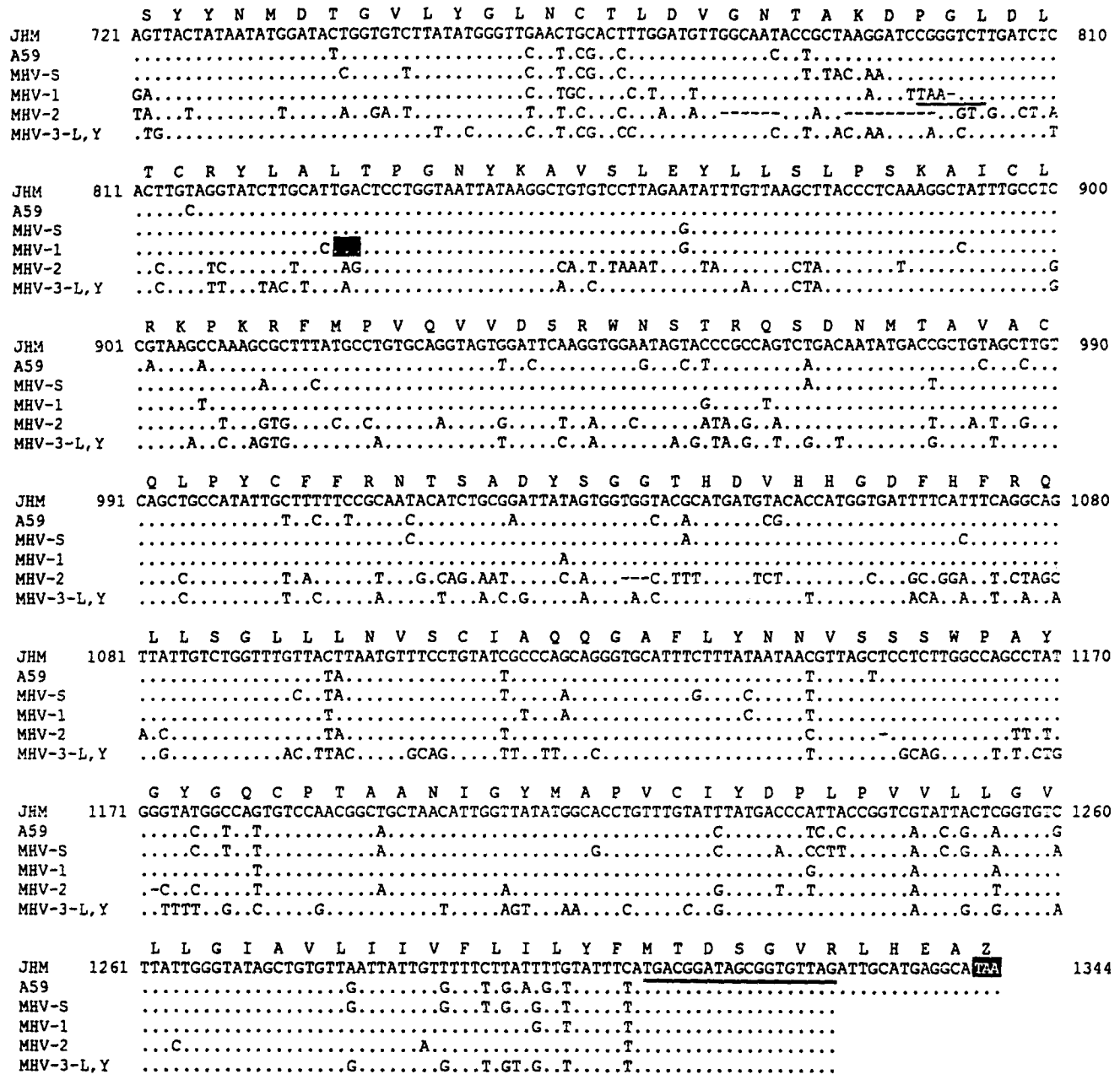

FIG. 3—*Continued*

than mRNA 2-1 and larger than mRNA 3 (Fig. 2). By performing PCR with a pair of oligonucleotides, one identical with the 5′-end of the leader RNA and the other complementary to the 3′-end region of the HE gene, the initiation site of this novel mRNA was determined and compared with the corresponding genomic sequence (Fig. 6). As a result of mutations and a deletion in this region, the genomic sequence from nucleotides 794 to 804 (ATCCGGGTCTT) was converted to ATTTAA-TCTT, which resembled the consensus transcription initiation signal. Comparison of the genomic sequence with the sequence of the novel mRNA indicated that this was indeed the transcriptional initiation point for the mRNA species (Fig. 6). Thus, the specific expression of the novel mRNA in MHV-1 was most likely caused by the interaction of the three copies of UCUAA in the leader with the newly acquired intergenic site in the middle of the HE gene.

DISCUSSION

The results presented in this paper show that the functional integrity of the HE gene is not well maintained in murine coronaviruses. Of the MHV strains

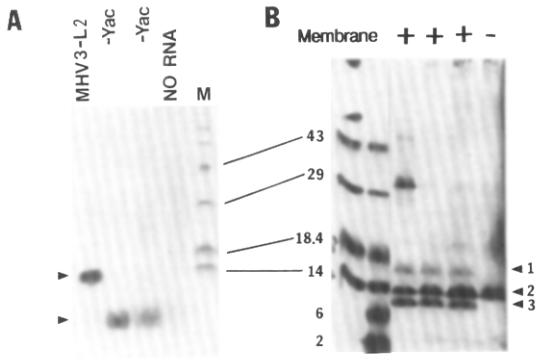

FIG. 4. SDS–PAGE analysis of *in vitro* translation products of the HE gene of MHV-3-L (L2) and MHV-3-Y (Yac). A. RNAs transcribed *in vitro* from cDNA clones were translated in a wheat germ extract system in the presence of [^{35}S]methionine (NEN Du pont). The arrowheads indicate the translation products from MHV-3-L and two different clones of MHV-3-Y. Lane M represents ^{14}C-labeled size markers (Gibco, BRL). "No RNA" lane was *in vitro* translation with no RNA. B. MHV-3-L RNA was translated in a rabbit reticulocyte lysate system with [^{35}S]methionine. The two left lanes are molecular size markers. Membrane "+" lanes represent products translated in the presence of the canine pancreatic microsomal membrane. The membrane "–" lane was translated without membrane. Products were immunoprecipitated with either monoclonal anti-HE antibody (the third lane from left) or anti-JHM(2) polyclonal antibody (the forth and fifth lanes from left, which represented different *in vitro* translation reactions). The specific precipitates are indicated by arrowheads with a number; 1, glycosylated product; 2, native translation product; 3, core product from which the signal peptide had been cleaved off. The band which migrates around 30 kDa appears to be a nonspecific precipitate, since it was detected in only one out of three different immunoprecipitation reactions.

examined, only two viruses (JHM and MHV-S) made an intact HE protein. Other viruses had a defective gene or were defective in transcription such that no HE protein was synthesized. The failure to synthesize an HE protein was due to several different reasons in different MHV strains: (1) The mRNA for the HE protein (mRNA 2-1) was not synthesized because the number of copies of the UCUAA pentanucleotide in the leader sequence was not optimum for transcription (e.g., JHM(3) and MHV-1, which contain three UCUAA copies), or the consensus intergenic sequence (e.g., A59) was mutated. (2) The mRNA was synthesized but due to sequence alterations resulting from mutations, deletions or insertions, the usual initiation codon of the HE ORF was lost, or the reading frame closed prematurely (e.g., A59, MHV-2, and MHV-3-L and MHV-3-Y). These findings are in striking contrast to the other three MHV structural protein genes, S, M, and N, which, despite some degree of sequence diversity, are consistently expressed and the ORFs are well-maintained. These results suggest that there is no selection pressure to preserve the complete HE ORF, and consequently, this gene has drifted into a pseudogene. Furthermore,

TABLE 1

PUTATIVE INTERGENIC SITE: LEADER FUSION SITE OF mRNA 2-1

Leader	(UCUAA)UCUAAUCUAAACUU		
HE gene	JHM		5′-uaa**UAAACUU**auua
	A59		-uaa**UAAGCUU**auua
	MHV-S		-uaa**UAAACUU**auua
	HMV-1		-uaa**UAAACUU**auua
	MHV-2		-uaa**UAAACUU**auua
	MHV-3-L		-AaUA**AAUCUAA**uAa
	MHV-3-Y		-AaUA**AAUCUAA**uAa
Gene 3 (JHM)			agca**UAAUCUAAAC**aug
Gene 7 (JHM)			**AAUCUAAUCUAAACUU**u
Novel mRNA	MHV-1	795	**UUUAAUCU**uga
	(JHM	795	uCCGGGucuuga)

Note. The leader fusion sites of mRNA 2-1 of various MHV strains were obtained from PCR products of the 5′ region of the mRNA, and are compared with the corresponding regions of genomic sequences. Boldface capital letters represent common sequences between leader RNA and the intergenic sequence of the genome, and thus are likely to be the leader fusion site. Divergent nucleotides are shown in regular capital letters. Other nucleotides are shown in lower case. As a comparison, the gene 3 and gene 7 intergenic site and the novel mRNA initiaion site within the HE gene of MHV-1 are shown. The sequence in parenthesis represents the corresponding JHM sequence in the HE gene.

since some of the MHV strains examined in this study did not express HE, this protein is not essential for viral replication. This conclusion is consistent with the hypothesis that this gene might have been derived from influenza C virus by a fortuitous RNA-RNA recombination event (Luytjes *et al.*, 1988). Nevertheless, the JHM and MHV-S viruses have maintained this gene despite extensive passages in tissue culture and animals.

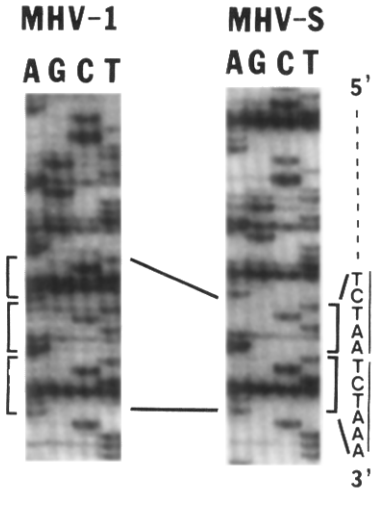

FIG. 5. RNA sequencing of the 3′-end region of the leader RNA of MHV-1 and MHV-S. Top, 5′-end; bottom, 3′-end. The brackets indicate the UCUAA pentanucleotide.

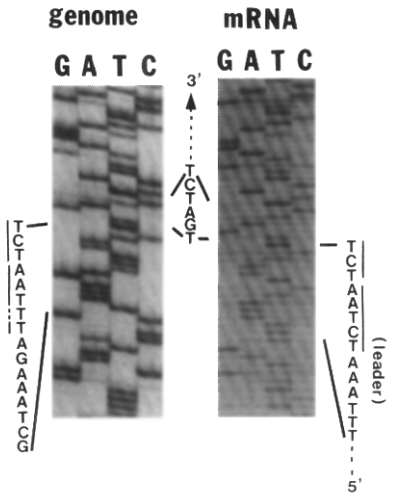

FIG. 6. DNA sequencing of the junction region between the leader and the body of the novel mRNA of MHV-1. The sequences are extended from the bottom to the top, in the 5' to 3' direction. The sidelined genomic sequence (left side) corresponds to the sidelined UCUAA sequence in the leader (right side). The sequence indicated in the middle shows the sequence which is homologous between the mRNA and the genome in the HE gene.

Also, the JHM virus with an expressed HE protein was naturally selected during JHM passages in tissue culture (Makino and Lai, 1989a). Thus, the presence of HE protein may have provided selective advantages under certain conditions. Interestingly, cold-sensitive mutants isolated from a persistent MHV-S infection of Kirsten murine sarcoma virus-transformed BALB/C cells failed to make HE protein (Yoshikura and Tejima, 1981). Also, Morris *et al.* (1989) have isolated JHM variants from Wistar Furth rats with a JHM-induced demyelinating disease. Besides a change in the S protein, these variants had a large deletion in the HE gene, resulting in the loss of HE protein expression (La Monica *et al.*, 1991). These data further suggest that under certain conditions, the HE protein may provide some selection advantage.

The HE protein also is present in some other coronaviruses which have a hemagglutinin activity including BCV and human coronavirus OC43, (Vlasak *et al.*, 1988b). The HE protein of BCV also exhibits an acetylesterase activity similar to the receptor destroying activity of influenza C virus (Vlasak *et al.*, 1988b). A similar activity has been demonstrated for the HE protein of MHV (Yokomori *et al.*, 1989). Data have been obtained which suggest that the HE protein is required for BCV infectivity. HE protein was the target of neutralizing monoclonal antibodies, which also inhibited hemagglutinin and esterase activities (Deregt and Babiuk, 1987; Parker *et al.*, 1990). In addition, treatment with diisopropyl fluorophosphate (DFP), an inhibitor of esterase, was found to inhibit virus replication (Vlasak *et al.*, 1988a). Similar requirements have not been demonstrated for MHV. It should be noted, however, that the presence of the HE protein does alter some biological properties of the virus. For example, in cultured cells, JHM(2) became dominant over JHM(3) after serial undiluted passages (Makino and Lai, 1989a). In mice, JHM(2) showed greater virulence and neurotropism than JHM(3) (Yokomori, unpublished data). Finally, passive immunization with monoclonal antibodies against HE protein protected mice from JHM(2) infection (Yokomori, unpublished data). These data suggested that HE may play a role in neuronal infection.

Although MHV-3 isolates examined retained a truncated HE ORF, and a protein product could be synthesized and properly processed in an *in vitro* translation system, no HE-related proteins could be detected in infected cells. The failure to detect these truncated HE proteins *in vivo* was not readily apparent. Perhaps the predicted initiating AUG was not in an ideal context in infected cells, although the sequence around this AUG agrees with the optimum translation context described by Kozak (1987). Also, the truncated protein products may degrade rapidly; however, even short pulse-labeling (15 min) did not reveal any products (Fig. 1). Additionally, because the truncated ORFs retained the putative esterase domain and signal sequence but lacked the C-terminal membrane-anchoring domain, we have investigated the possibility that the truncated HE protein may have been released into the media. No HE-specific esterase activity was detected in supernatants from infected cells (data not shown). Thus, the reason that no truncated HE protein was detected in these MHV strains remains unclear.

Another interesting characteristic of the HE gene is its transcriptional control by the copy number of UCUAA pentanucleotide sequence in the leader RNA. Both this report and previous publications (Makino and Lai, 1989a; Yokomori *et al.*, 1989) indicate that only JHM variants with two copies of UCUAA in the leader expresses mRNA 2-1 efficiently, whereas JHM variants with three copies of UCUAA transcribe only small amounts of this mRNA (Makino and Lai, 1989; Shieh *et al.*, 1989). This mechanism of transcriptional control also was confirmed with a variant of JHM, JHM-DL. Furthermore, we found that other MHV strains also utilized the same transcriptional regulation. Strains MHV-S, MHV-2, and MHV-3 had two UCUAA copies and expressed a large amount of mRNA 2-1, whereas MHV-1 possessed three copies and did not express mRNA 2-1, even though it had the same intergenic sequence preceding the gene as other MHVs. In addition, MHV-1 synthesized a smaller mRNA from a novel downstream site which had a sequence similar to the

consensus transcriptional initiation signal (UUUAAUCUU vs UCUAAUCUA, respectively). These results strongly suggest that MHV RNA transcription results from an interaction between the leader RNA and the intergenic sequence preceding each gene, although the precise mechanism for this transcriptional control is not yet clear. The differential control of HE gene transcription by the leader RNA may further our understanding of the mechanism of MHV transcription.

ACKNOWLEDGMENTS

We thank Dr. Shinji Makino for providing passaged JHM-DL, Kenneth Lu for help in DNA sequencing, and Drs. Tom MacNaughton and John Polo for critical comments and editorial assistance. This work was supported by US Public Health Services Research Grants AI 19244 and NS 18146 from the National Institutes of Health. K.Y. and L.R.B. are supported in part by Feiger fellowships from Norris Cancer Center, University of Southern California. M.M.C.L. is an investigator of Howard Hughes Medical Institute.

REFERENCES

BANNER, L. R., KECK, J. G., and LAI, M. M. C. (1990). A clustering of RNA recombination sites adjacent to a hypervariable region of the peplomer gene of murine coronavirus. *Virology* **175,** 548–555.

COLLINS, A. R., KNOBLER, R. L., POWELL, H., and BUCHMEIER, M. J. (1982). Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell–cell fusion. *Virology* **119,** 358–371.

DEA, S., and TIJSSEN, P. (1988). Identification of the structural proteins of turkey enteric coronavirus. *Arch. Virol.* **99,** 173–186.

DEREGT, D., and BABIUK, L. A. (1987). Monoclonal antibodies to bovine coronavirus: Characteristics and topological mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* **161,** 410–420.

GARWES, D. J., and REYNOLDS, D. J. (1981). The polypeptide structure of canine coronavirus and its relationship to porcine transmissible gastroenteritis virus. *J. Gen. Virol.* **52,** 153–157.

HIRANO, N., FUJIWARA, K., HINO, S., and MATSUMOTO, M. (1974). Replication and plaque formation of mouse hepatitis virus (MHV-2) in mouse cell line DBT culture. *Arch. Gesamte Virusforsch.* **44,** 298–302.

HOGUE, B. G., and BRIAN, D. A. (1986). Structural proteins of human respiratory coronavirus OC43. *Virus Res.* **5,** 131–144.

HOGUE, B. G., KING, B., and BRIAN, D. A. (1984). Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. *J. Virol.* **51,** 384–388.

KING, B., POTTS, B. J., and BRIAN, D. A. (1985). Bovine coronavirus haemagglutinin protein. *Virus Res.* **2,** 53–59.

KOZAK, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15,** 8125–8148.

LA MONICA, N., BANNER, L. R., MORRIS, V. L., and LAI, M. M. C. (1991). Localization of extensive deletions in the structural genes of two neurotropic variants of murine coronavirus JHM. *Virology* **182,** 883–888.

LAI, M. M. C. (1990). Coronavirus: Organization, replication and expression of genome. *Annu. Rev. Microbiol.* **44,** 303–333.

LAI, M. M. C., BARIC, R. S., BRAYTON, P. R., and STOHLMAN, S. A. (1984). Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc. Natl. Acad. Sci. USA* **81,** 3626–3630.

LAI, M. M. C., BRAYTON, P. R., ARMEN, R. C., PATTON, C. D., PUGH, C., and STOHLMAN, S. A. (1981). Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. *J. Virol.* **39,** 823–834.

LAI, M. M. C., and STOHLMAN, S. A. (1981). Comparative analysis of RNA genome of mouse hepatitis virus. *J. Virol.* **38,** 661–670.

LEE, H.-J., SHIEH, C.-K., GORBALENYA, A. E., KOONIN, E. V., LA MONICA, N., TULER, J., BAGDZHADZHYAN, A., and LAI, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180,** 567–582.

LEIBOWITZ, J. L., WEISS, S. R., PAAVOLA, E., and BOND, C. W. (1982). Cell-free translation of murine coronavirus RNA. *J. Virol.* **43,** 905–913.

LEIBOWITZ, J. L., WILHELMSEN, K. C., and BOND, C. W. (1981). The virus specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. *Virology* **114,** 39–51.

LUYTJES, W., BREDENBEEK, P. J., NOTEN, A. F. H., HORZINEK, M. C., and SPAAN, W. J. M. (1988). Sequence of mouse hepatitis virus A59 mRNAs: Indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166,** 415–422.

MAKINO, S., and LAI, M. M. C. (1989a). Evolution of the 5′-end of genomic RNA of murine coronaviruses during passages in vitro. *Virology* **169,** 227–232.

MAKINO, S., and LAI, M. M. C. (1989b). High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *J. Virol.* **63,** 5285–5292.

MAKINO, S., SOE, L. H., SHIEH, C.-K., and LAI, M. M. C. (1988). Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNA. *J. Virol.* **62,** 3870–3873.

MAKINO, S., TAGUCHI, F., and FUJIWARA, K. (1984a). Defective interfering particles of mouse hepatitis virus. *Virology* **133,** 9–17.

MAKINO, S., TAGUCHI, F., HIRANO, N., and FUJIWARA, K. (1984b). Analysis of genomic and intracellular viral RNAs of small plaque mutants of mouse hepatitis virus, JHM strain. *Virology* **139,** 138–151.

MORRIS, V. L., TIESZER, C., MACKINNON, J., and PERCY, D. (1989). Characterization of coronavirus JHM variants isolated from Wistar Furth rats with a viral-induced demyelinating disease. *Virology* **169,** 127–136.

PACHUK, C. J., BREDENBEEK, P. J., ZOLTICK, P. W., SPAAN, W. J. M., and WEISS, S. R. (1989). Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus strain A59. *Virology* **171,** 141–148.

PARKER, M. D., COX, G. J., DEREGT, D., FITZPATRICK, D. R., and BABIUK, L. A. (1989). Cloning and in vitro expression of the gene for the E3 haemagglutinin glycoprotein of bovine coronavirus. *J. Gen Virol.* **70,** 155–164.

PARKER, M. D., YOO, D., and BABIUK, L. A. (1990). Expression and secretion of the bovine coronavirus hemagglutinin-esterase glycoprotein by insect cells infected with recombinant baculoviruses. *J. Virol.* **64,** 1625–1629.

SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

SHIEH, C.-K., LEE, H.-J., YOKOMORI, K., LA MONICA, N., MAKINO, S., and LAI, M. M. C. (1989). Identification of a new transcriptional initiation site and the corresponding functional gene 2b in the murine coronavirus RNA genome. *J. Virol.* **63,** 3729–3736.

SIDDELL, S. (1983). Coronavirus JHM: Coding assignments of subgenomic mRNAs. *J. Gen. Virol.* **64,** 113–125.

SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., VAN DER ZEIJST, B. A. M., and SIDDELL, S. G. (1983).

Coronavirus mRNA synthesis involves fusion of noncontiguous sequences. *EMBO J.* **2,** 1839–1844.

STERN, D. F., and SEFTON, B. M. (1982). Coronavirus proteins: Structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *J. Virol.* **44,** 804–812.

STOHLMAN, S. A., BARIC, R. S., NELSON, G. N., SOE, L. H., WELTER, L. M., and DEANS, R. J. (1988). Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* **62,** 4288–4295.

STOHLMAN, S. A., BRAYTON, P. R., FLEMING, J. O., WEINER, L. P., and LAI, M. M. C. (1982). Murine coronaviruses: Isolation and characterization of two plaque morphology variants of the JHM neurotropic strain. *J. Gen. Virol.* **63,** 265–275.

STURMAN, L. S., and HOLMES, K. V. (1983). The molecular biology of coronaviruses. *Adv. Virus Res.* **28,** 35–112.

STURMAN, L. S., HOLMES, K. V., and BEHNKE, J. (1980). Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* **33,** 449–462.

STURMAN, L. S., RICARD, C. S., and HOLMES, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: Activation of cell fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J. Virol.* **56,** 904–911.

VLASAK, R., LUYTJES, W., LEIDER, J., SPAAN, W., and PALESE, P. (1988a). The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. *J. Virol.* **62,** 4686–4690.

VLASAK, R., LUYTJES, W., SPAAN, W., and PALESE, P. (1988b). Human and bovine coronaviruses recognize sialic acid-containing receptor similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. USA* **85,** 4526–4529.

WEINER, L. P. (1973). Pathogenesis of demyelination induced by a mouse hepatitis virus (JHM virus). *Arch. Neurol.* **28,** 298–303.

YOKOMORI, K., LA MONICA, N., MAKINO, S., SHIEH, C.-K., and LAI, M. M. C. (1989). Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. *Virology* **173,** 683–691.

YOSHIKURA, H., and TEJIMA, S. (1981). Role of protease in mouse hepatitis virus-induced cell fusion. *Virology* **113,** 503–511.